GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 11:45:56 ; Search time 3959 Seconds
(without alignments)
16047.651 Million cell updates/sec

Title: US-09-745-506-74
Perfect score: 1553
Sequence: 1 GTGATTGTTATCTTGGTGCT.....TCTGTTTACTTAACATTCAA 1553

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		1529	98.5	1606	9	AB038949	AB038949 Homo sapi
2		1515	97.6	1574	6	BD158389	BD158389 Primer fo
3		1515	97.6	1574	9	AK023378	AK023378 Homo sapi
4		1487.6	95.8	1579	9	AF182416	AF182416 Homo sapi
5		1338.4	86.2	1387	9	AF060513	AF060513 Homo sapi
6		1336.6	86.1	1440	9	BC007654	BC007654 Homo sapi
7		1334	85.9	1425	9	AF283538	AF283538 Homo sapi
8		1232	79.3	1353	9	HSM805522	AL834430 Homo sapi
9		1053	67.8	1053	6	AX119075	AX119075 Sequence
10		868	55.9	1836	10	AF284439	AF284439 Mus muscu
11		725	46.7	796	6	BD149184	BD149184 Primer fo
12	c	708.4	45.6	55374	9	AL645474	AL645474 Human DNA
13	c	417	26.9	170586	2	AC037455	AC037455 Homo sapi
14	c	417	26.9	190508	9	AC005037	AC005037 Homo sapi
15		414.4	26.7	198250	2	AC093681	AC093681 Homo sapi
16		329.6	21.2	149819	2	AC121091	AC121091 Mus muscu
17		329.6	21.2	234976	2	AC118698	AC118698 Mus muscu
18		325.4	21.0	231600	2	AC130779	AC130779 Rattus no
19		224.8	14.5	249	6	BD049005	BD049005 Sequence
20		223.2	14.4	1328	3	AK114307	AK114307 Ciona int
21	c	145.2	9.3	281017	2	AC123462	AC123462 Rattus no
22	c	141	9.1	198250	2	AC093681	AC093681 Homo sapi
23	c	134.6	8.7	155127	2	AL929495	AL929495 Danio rer
24	c	117.6	7.6	231600	2	AC130779	AC130779 Rattus no
25		108.4	7.0	1036	3	AY069556	AY069556 Drosophil
26		108.4	7.0	70667	2	AC017306	AC017306 Drosophil
27		108.4	7.0	165839	3	AC092398	AC092398 Drosophil
28		108.4	7.0	273414	3	AE003650	AE003650 Drosophil
29		108.4	7.0	299935	3	DROSADH09	AE003415 Drosophil
30	c	97.8	6.3	157118	9	AC069286	AC069286 Homo sapi
31		90	5.8	339	6	BD040994	BD040994 Sequence
32	c	53.2	3.4	12320	1	AE013129	AE013129 Thermoana
33		53	3.4	5463	6	AR227150	AR227150 Sequence
34		53	3.4	28324	8	SPCC126	AL034490 S.pombe c
35	c	49.6	3.2	218470	1	BSUB0013	Z99116 Bacillus su
36		49.6	3.2	282700	1	BACJH642	D84432 Bacillus su
37		49.6	3.2	299550	1	AP001511	AP001511 Bacillus
38		47.6	3.1	805	6	AX434353	AX434353 Sequence
39	c	46.6	3.0	1187	8	SCYGL221C	Z72743 S.cerevisia
40		45.8	2.9	348250	1	AP003592	AP003592 Nostoc sp
41		45.6	2.9	302173	1	AE016951	AE016951 Enterococ
42	c	43.6	2.8	300800	1	SCO939112	AL939112 Streptomy
43		43.4	2.8	209604	2	AC101593	AC101593 Mus muscu
44	c	43.2	2.8	2000	6	AX655393	AX655393 Sequence
45	c	43.2	2.8	299050	1	BX251411	BX251411 Tropherym

ALIGNMENTS

RESULT 1
AB038949
LOCUS AB038949 1606 bp mRNA linear PRI 23-FEB-2001
DEFINITION Homo sapiens ALS2CR1 mRNA for amyotrophic lateral sclerosis 2, candidate 1, complete cds.
ACCESSION AB038949
VERSION AB038949.1 GI:12862477
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Hadano,S., Yanagisawa,Y., Skaug,J., Fichter,K., Nasir,J., Martindale,D., Koop,B.F., Scherer,S.W., Nicholson,D.W.,

```
            Rouleau,G.A., Ikeda,J.-E. and Hayden,M.R.
  TITLE     Cloning and characterization of three novel genes, ALS2CR1,
            ALS2CR2, and ALS2CR3, in the juvenile amyotrophic lateral sclerosis
            (ALS2) critical region at chromosome 2q33-q34: candidate genes for
            ALS2
  JOURNAL   Genomics 71 (2), 200-213 (2001)
  MEDLINE   21100893
   PUBMED   11161814
REFERENCE   2  (bases 1 to 1606)
  AUTHORS   Hadano,S., Ikeda,J. and Hayden,M.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-FEB-2000) Shinji Hadano, Tokai University, The
            Institute of Medical Sciences; Bohseidai, Isehara, Kanagawa
            259-1193, Japan (E-mail:shinji@nga.med.u-tokai.ac.jp,
            Tel:+81-463-91-5095, Fax:+81-463-91-4993)
FEATURES             Location/Qualifiers
     source          1. .1606
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2q33"
                     /note="IMAGE:32584~UniGene ID: Hs.21943"
     gene            1. .1606
                     /gene="ALS2CR1"
     CDS             286. .1338
                     /gene="ALS2CR1"
                     /note="candidate 1"
                     /codon_start=1
                     /product="amyotrophic lateral sclerosis 2"
                     /protein_id="BAB32499.1"
                     /db_xref="GI:12862478"
                     /translation="MDLKALLSSLNDFASLSFAESWDNVGLLVEPSPPHTVNTLFLTN
                     DLTEEVMEEVLQKKADLILSYHPPIFRPMKRITWNTWKERLVIRALENRVGIYSPHTA
                     YDAAPQGVNNWLAKGLGACTSRPIHPSKAPNYPTEGNHRVEFNVNYTQDLDKVMSAVK
                     GIDGVSVTSFSARTGNEEQTRINLNCTQKALMQVVDFLSRNKQLYQKTEILSLEKPLL
                     LHTGMGRLCTLDESVSLATMIDRIKRHLKLSHIRLALGVGRTLESQVKVVALCAGSGS
                     SVLQGVEADLYLTGEMSHHDTLDAASQGINVILCEHSNTERGFLSDLRDMLDSHLENK
                     INIILSETDRDPLQVV"
     polyA_signal    1566. .1571
                     /gene="ALS2CR1"
BASE COUNT      438 a    365 c    375 g    428 t
ORIGIN

  Query Match           98.5%;  Score 1529;  DB 9;  Length 1606;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1551;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy          1 GTGATTGTTATCTTGGTGCTGCAGAGGACAGCAGAAGAGGAGATTGGGTCAGAAAACTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         42 GTGATTGTTATCTTGGTGCTGCAGAGGACAGCAGAAGAGGAGATTGGGTCAGAAAACTGC 101

Qy         61 CCTGCCGCACCAGAGCACAGCGCACTAGTGGGACAGGGGTCCTGACTCAGACTTAACTGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        102 CCTGCCGCACCAGAGCACAGCGCACTAGTGGGACAGGGGTCCTGACTCAGACTTAACTGG 161

Qy        121 CTGTGTCTCGTGGTTTTTCACTGTCCTGGAAAAGGCCTGAAGTGGCACTGAAATGAGGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 CTGTGTCTCGTGGTTTTTCACTGTCCTGGAAAAGGCCTGAAGTGGCACTGAAATGAGGCA 221

Qy        181 TAGATGAGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 TAGATGAGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTC 281

Qy        241 CTTCATGGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 CTTCATGGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGC 341

Qy        301 TGAGAGTTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 TGAGAGTTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATAC 401

Qy        361 ACTCTTCCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 ACTCTTCCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGC 461

Qy        421 AGACCTCATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 AGACCTCATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAA 521

Qy        481 CACATGGAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 CACATGGAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCC 581

Qy        541 TCATACAGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 TCATACAGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGG 641

Qy        601 AGCTTGTACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 AGCTTGTACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAA 701

Qy        661 CCACCGAGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 CCACCGAGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGT 761

Qy        721 GAAAGGAATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGAACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        762 GAAAGGAATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGAACA 821

Qy        781 AACACGGATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCTTTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        822 AACACGGATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCTTTC 881

Qy        841 CCGGAACAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCTTCT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        882 CCGGAACAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCTTCT 941

Qy        901 ACATACTGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCATGAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        942 ACATACTGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCATGAT 1001

Qy        961 TGATCGAATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGGGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1002 TGATCGAATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGGGAG 1061

Qy       1021 AACCTTAGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1062 AACCTTAGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCT 1121

Qy       1081 GCAGGGTGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTTGGA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1122 GCAGGGTGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTTGGA 1181

Qy       1141 TGCTGCTTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGGCTT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1182 TGCTGCTTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGGCTT 1241

Qy       1201 TCTTTCTGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTATCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1242 TCTTTCTGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTATCCT 1301

Qy       1261 ATCAGAGACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAACAC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1302 ATCAGAGACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAACAC 1361

Qy       1321 ATTCCTACAAATCAGCTGGATGCCCAACTTAAATTTGTAACATGAGTCAGTGGGACTGGT 1380
              ||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1362 ATT-CTACAAATCAGCTGGATG-CCAACTTAAATTTGTAACATGAGTCAGTGGGACTGGT 1419

Qy       1381 GTGCTTCCAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTTATTCACC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1420 GTGCTTCCAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTTATTCACC 1479

Qy       1441 AAATGTTCTATCGCTCGTAAGGTAAAACTGTAATATAACTACCATATTAAATAACAAATG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1480 AAATGTTCTATCGCTCGTAAGGTAAAACTGTAATATAACTACCATATTAAATAACAAATG 1539

Qy       1501 TTCATTATAAACTCTAGGAAAGATTGAATAAAATCTGTTTACTTAACATTCAA 1553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1540 TTCATTATAAACTCTAGGAAAGATTGAATAAAATCTGTTTACTTAACATTCAA 1592


RESULT 2
BD158389
LOCUS       BD158389               1574 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD158389
VERSION     BD158389.1  GI:27864147
KEYWORDS    JP 2002191363-A/13232.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1574)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 13232 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002191363-A/13232
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                     Location/Qualifiers
            FT   CDS             (271). .(1320).
FEATURES             Location/Qualifiers
     source          1. .1574
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      420 a    361 c    372 g    421 t
ORIGIN

  Query Match           97.6%;  Score 1515;  DB 6;  Length 1574;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 1548;  Conservative    0;  Mismatches    0;  Indels    3;  Gaps    3;

Qy          1 GTGATTGTTATCTTGGTGCTGCAGAGGACAGCAGAAGAGGAGATTGGGTCAGAAAACTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         27 GTGATTGTTATCTTGGTGCTGCAGAGGACAGCAGAAGAGGAGATTGGGTCAGAAAACTGC 86

Qy         61 CCTGCCGCACCAGAGCACAGCGCACTAGTGGGACAGGGGTCCTGACTCAGACTTAACTGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 CCTGCCGCACCAGAGCACAGCGCACTAGTGGGACAGGGGTCCTGACTCAGACTTAACTGG 146

Qy        121 CTGTGTCTCGTGGTTTTTCACTGTCCTGGAAAAGGCCTGAAGTGGCACTGAAATGAGGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CTGTGTCTCGTGGTTTTTCACTGTCCTGGAAAAGGCCTGAAGTGGCACTGAAATGAGGCA 206

Qy        181 TAGATGAGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TAGATGAGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTC 266

Qy        241 CTTCATGGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CTTCATGGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGC 326

Qy        301 TGAGAGTTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        327 TGAGAGTTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATAC 386

Qy        361 ACTCTTCCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ACTCTTCCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGC 446

Qy        421 AGACCTCATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 AGACCTCATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAA 506

Qy        481 CACATGGAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CACATGGAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCC 566

Qy        541 TCATACAGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCATACAGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGG 626

Qy        601 AGCTTGTACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 AGCTTGTACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAA 686

Qy        661 CCACCGAGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 CCACCGAGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGT 746

Qy        721 GAAAGGAATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGAACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 GAAAGGAATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGAACA 806

Qy        781 AACACGGATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCTTTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 AACACGGATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCTTTC 866

Qy        841 CCGGAACAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCTTCT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 CCGGAACAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCTTCT 926

Qy        901 ACATACTGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCATGAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 ACATACTGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCATGAT 986

Qy        961 TGATCGAATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGGGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 TGATCGAATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGGGAG 1046

Qy       1021 AACCTTAGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 AACCTTAGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCT 1106

Qy       1081 GCAGGGTGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTTGGA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 GCAGGGTGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTTGGA 1166

Qy       1141 TGCTGCTTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGGCTT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 TGCTGCTTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGGCTT 1226

Qy       1201 TCTTTCTGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTATCCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 TCTTTCTGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTATCCT 1286

Qy       1261 ATCAGAGACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAACAC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 ATCAGAGACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAACAC 1346

Qy       1321 ATTCCTACAAATCAGCTGGATGCCCAACTTAAATTTGTAACATGAGTCAGTGGGACTGGT 1380
              ||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1347 ATT-CTACAAATCAGCTGGATG-CCAACTTAAATTTGTAACATGAGTCAGTGGGACTGGT 1404

Qy       1381 GTGCTTCCAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTTATTCACC 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1405 GTGCTTCCAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTTATTCACC 1464
```

```
Qy       1441 AAATGTTCTATCGCTCGTAAGGTAAAACTGTAATATAACTACCATATTAAATAACAAATG 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1465 AAATGTTCTATCGCTCGTAAGGTAAAACTGTAATATAACTACCATATTAAATAACAAATG 1524

Qy       1501 TTCATTATAAACTCTAGGAAAGATTGAATAAAATCTGTTTACTTAACATTC 1551
             |||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1525 TTCATTATAAA-TCTAGGAAAGATTGAATAAAATCTGTTTACTTAACATTC 1574


RESULT 3
AK023378
LOCUS       AK023378               1574 bp    mRNA    linear   PRI 01-AUG-2002
DEFINITION  Homo sapiens cDNA FLJ13316 fis, clone OVARC1001555, weakly similar
            to NGG1-INTERACTING FACTOR 3.
ACCESSION   AK023378
VERSION     AK023378.1  GI:10435289
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
            Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
            Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
            Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
            Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
            Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1574)
  AUTHORS   Isogai,T. and Otsuki,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing and clone selection:
            Helix Research Institute (supported by Japan Key Technology Center
            etc.) and Department of Virology, Institute of Medical Science,
            University of Tokyo.
FEATURES             Location/Qualifiers
     source          1. .1574
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="OVARC1001555"
                     /tissue_type="ovary, tumor tissue"
                     /clone_lib="OVARC1"
                     /note="cloning vector: pME18SFL3"
     CDS             271. .1323
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAB14551.1"
                     /db_xref="GI:10435290"
                     /translation="MDLKALLSSLNDFASLSFAESWDNVGLLVEPSPPHTVNTLFLTN
                     DLTEEVMEEVLQKKADLILSYHPPIFRPMKRITWNTWKERLVIRALENRVGIYSPHTA
                     YDAAPQGVNNWLAKGLGACTSRPIHPSKAPNYPTEGNHRVEFNVNYTQDLDKVMSAVK
                     GIDGVSVTSFSARTGNEEQTRINLNCTQKALMQVVDFLSRNKQLYQKTEILSLEKPLL
                     LHTGMGRLCTLDESVSLATMIDRIKRHLKLSHIRLALGVGRTLESQVKVVALCAGSGS
                     SVLQGVEADLYLTGEMSHHDTLDAASQGINVILCEHSNTERGFLSDLRDMLDSHLENK
                     INIILSETDRDPLQVV"
BASE COUNT      420 a    361 c    372 g    421 t
ORIGIN

  Query Match             97.6%;  Score 1515;  DB 9;  Length 1574;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1548;  Conservative    0;  Mismatches    0;  Indels    3;  Gaps    3;
```

```
Qy          1 GTGATTGTTATCTTGGTGCTGCAGAGGACAGCAGAAGAGGAGATTGGGTCAGAAAACTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         27 GTGATTGTTATCTTGGTGCTGCAGAGGACAGCAGAAGAGGAGATTGGGTCAGAAAACTGC 86

Qy         61 CCTGCCGCACCAGAGCACAGCGCACTAGTGGGACAGGGGTCCTGACTCAGACTTAACTGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 CCTGCCGCACCAGAGCACAGCGCACTAGTGGGACAGGGGTCCTGACTCAGACTTAACTGG 146

Qy        121 CTGTGTCTCGTGGTTTTTCACTGTCCTGGAAAAGGCCTGAAGTGGCACTGAAATGAGGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CTGTGTCTCGTGGTTTTTCACTGTCCTGGAAAAGGCCTGAAGTGGCACTGAAATGAGGCA 206

Qy        181 TAGATGAGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TAGATGAGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTC 266

Qy        241 CTTCATGGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CTTCATGGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGC 326

Qy        301 TGAGAGTTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 TGAGAGTTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATAC 386

Qy        361 ACTCTTCCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ACTCTTCCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGC 446

Qy        421 AGACCTCATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 AGACCTCATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAA 506

Qy        481 CACATGGAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CACATGGAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCC 566

Qy        541 TCATACAGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCATACAGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGG 626

Qy        601 AGCTTGTACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 AGCTTGTACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAA 686

Qy        661 CCACCGAGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 CCACCGAGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGT 746

Qy        721 GAAAGGAATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGAACA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 GAAAGGAATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGAACA 806

Qy        781 AACACGGATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCTTTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 AACACGGATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCTTTC 866

Qy        841 CCGGAACAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCTTCT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 CCGGAACAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCTTCT 926

Qy        901 ACATACTGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCATGAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 ACATACTGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCATGAT 986

Qy        961 TGATCGAATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGGGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        987 TGATCGAATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGGGAG 1046

Qy       1021 AACCTTAGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1047 AACCTTAGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCT 1106

Qy       1081 GCAGGGTGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTTGGA 1140
```

```
Db      1107 GCAGGGTGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTTGGA 1166

Qy      1141 TGCTGCTTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGGCTT 1200
Db      1167 TGCTGCTTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGGCTT 1226

Qy      1201 TCTTTCTGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTATCCT 1260
Db      1227 TCTTTCTGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTATCCT 1286

Qy      1261 ATCAGAGACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAACAC 1320
Db      1287 ATCAGAGACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAACAC 1346

Qy      1321 ATTCCTACAAATCAGCTGGATGCCCAACTTAAATTTGTAACATGAGTCAGTGGGACTGGT 1380
Db      1347 ATT-CTACAAATCAGCTGGATG-CCAACTTAAATTTGTAACATGAGTCAGTGGGACTGGT 1404

Qy      1381 GTGCTTCCAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTTATTCACC 1440
Db      1405 GTGCTTCCAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTTATTCACC 1464

Qy      1441 AAATGTTCTATCGCTCGTAAGGTAAAACTGTAATATAACTACCATATTAAATAACAAATG 1500
Db      1465 AAATGTTCTATCGCTCGTAAGGTAAAACTGTAATATAACTACCATATTAAATAACAAATG 1524

Qy      1501 TTCATTATAAACTCTAGGAAAGATTGAATAAAATCTGTTTACTTAACATTC 1551
Db      1525 TTCATTATAAA-TCTAGGAAAGATTGAATAAAATCTGTTTACTTAACATTC 1574
```

```
RESULT 4
AF182416
LOCUS       AF182416               1579 bp    mRNA    linear   PRI 20-SEP-2000
DEFINITION  Homo sapiens MDS015 (MDS015) mRNA, complete cds.
ACCESSION   AF182416
VERSION     AF182416.1  GI:10197631
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1579)
  AUTHORS   Huang,C., Qian,B., Tu,Y., Gu,W., Wang,Y., Han,Z. and Chen,Z.
  TITLE     Novel genes expressed in hematopoietic stem/progenitor cells from
            Myelodysplastic Syndromes patient
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1579)
  AUTHORS   Huang,C., Qian,B., Tu,Y., Gu,W., Wang,Y., Han,Z. and Chen,Z.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-SEP-1999) Chinese National Human Genome Center at
            Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
            Shanghai 201203, People's Republic of China
FEATURES             Location/Qualifiers
     source          1. .1579
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /cell_type="hematopoietic stem/progenitor cells"
                     /note="isolated from a Myelodysplastic Syndrome patient"
     gene            1. .1579
                     /gene="MDS015"
     CDS             244. .1299
                     /gene="MDS015"
                     /codon_start=1
                     /product="MDS015"
                     /protein_id="AAG14952.1"
                     /db_xref="GI:10197632"
                     /translation="MDLKALLSSLNDFASLSFAESWDNVGLLVEPSPPHTVNTLFLTN
                     DLTEEVMEEVLQKKADLILSYHPPIFRPMKRINLEHMGRERLVIRALENRVGIYSPHT
                     AYDAAPQGVNNWLAKGLGACTSRPIHPSKAPNYPTEGNHRVEINVNYTQDLDKVMSAV
                     KGIDGVSVTSFSARTGNEEQTRINLNCTQKALMQVVDFLSRNKQLYQKTEILSLEKPL
                     LLHTGMGRLCTLDESVSLATMIDRIKRHLKLSHIRLALGVGRTLESQVKVVALCAGSG
                     SSVLQGVEADLYLTGEMSHHDTLDAASQGINVILCEHSNTERGFLSDLRDMLDSHLEN
                     KINIILSETDRDPLQVV"
BASE COUNT      449 a    354 c    360 g    416 t
ORIGIN
```

```
  Query Match             95.8%;  Score 1487.6;  DB 9;  Length 1579;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1546;  Conservative    0;  Mismatches    4;  Indels    6;  Gaps    5;

Qy          1 GTGATTGTTATCTTGGTGCTGCAGAGGACAGCAGAAGAGGAGATTGGGTCAGAAAACTGC 60
Db          1 GTGATTGTTATCTTGGTGCTGCAGAGGACAGCAGAAGAGGAGATTGGGTCAGAAAACTGC 60

Qy         61 CCTGCCGCACCAGAGCACAGCGCACTAGTGGGACAGGGGTCCTGACTCAGACTTAACTGG 120
Db         61 CCTGCCGCACCAGAGCACAGCGCACTAGTGGGACAGGGGTCCTGACTCAGACTTAACTGG 120

Qy        121 CTGTGTCTCGTGGTTTTTCACTGTCCTGGAAAAGGCCTGAAGTGGCACTGAAATGAGGCA 180
Db        121 CTGTGT-TCGTGGTTTTTCACTGTCCTGGAAAAGGCCTGAAGTGGCACTGAAATGAGGCA 179

Qy        181 TAGATGAGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTC 240
Db        180 TAGATGAGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTC 239

Qy        241 CTTCATGGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGC 300
Db        240 CTTCATGGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGC 299

Qy        301 TGAGAGTTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATAC 360
Db        300 TGAGAGTTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATAC 359

Qy        361 ACTCTTCCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGC 420
Db        360 ACTCTTCCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGC 419

Qy        421 AGACCTCATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCA-TAACCTGGA 479
Db        420 AGACCTCATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATTAACCTGGA 479

Qy        480 ACACATGG--AAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTC 537
Db        480 ACACATGGGAAGGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTC 539

Qy        538 TCCTCATACAGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCT 597
Db        540 TCCTCATACAGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCT 599

Qy        598 TGGAGCTTGTACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGG 657
Db        600 TGGAGCTTGTACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGG 659

Qy        658 AAACCACCGAGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGC 717
Db        660 AAACCACCGAGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGC 719

Qy        718 AGTGAAAGGAATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGA 777
Db        720 AGTGAAAGGAATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGA 779

Qy        778 ACAAACACGGATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCT 837
Db        780 ACAAACACGGATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCT 839

Qy        838 TTCCCGGAACAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCT 897
Db        840 TTCCCGGAACAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCT 899

Qy        898 TCTACATACTGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCAT 957
Db        900 TCTACATACTGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCAT 959
```

```
Qy        958 GATTGATCGAATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGG 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 GATTGATCGAATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGG 1019

Qy       1018 GAGAACCTTAGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGT 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 GAGAACCTTAGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGT 1079

Qy       1078 TCTGCAGGGTGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTT 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 TCTGCAGGGTGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTT 1139

Qy       1138 GGATGCTGCTTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGG 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 GGATGCTGCTTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGG 1199

Qy       1198 CTTTCTTTCTGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTAT 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 CTTTCTTTCTGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTAT 1259

Qy       1258 CCTATCAGAGACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAA 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 CCTATCAGAGACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAA 1319

Qy       1318 CACATTCCTACAAATCAGCTGGATGCCCAACTTAAATTTGTAACATGAGTCAGTGGGACT 1377
              |||||| ||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db       1320 CACATT-CTACAAATCAGCTGGATG-CCAACTTAAATTTGTAACATGAGTCAGTGGGACT 1377

Qy       1378 GGTGTGCTTCCAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTTATTC 1437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1378 GGTGTGCTTCCAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTTATTC 1437

Qy       1438 ACCAAATGTTCTATCGCTCGTAAGGTAAAACTGTAATATAACTACCATATTAAATAACAA 1497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1438 ACCAAATGTTCTATCGCTCGTAAGGTAAAACTGTAATATAACTACCATATTAAATAACAA 1497

Qy       1498 ATGTTCATTATAAACTCTAGGAAAGATTGAATAAAATCTGTTTACTTAACATTCAA 1553
              ||||||||||||||||||||||||||||||||||||||||||||||||||   ||
Db       1498 ATGTTCATTATAAACTCTAGGAAAGATTGAATAAAATCTGTTTACTTAACAAAAAA 1553


RESULT 5
AF060513
LOCUS       AF060513                1387 bp    mRNA    linear   PRI 02-JAN-2001
DEFINITION  Homo sapiens clone 016d06 My018 protein mRNA, complete cds.
ACCESSION   AF060513
VERSION     AF060513.1  GI:12001975
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1387)
  AUTHORS   Mao,Y.M., Xie,Y., Huang,X.Y., Ying,K. and Dai,J.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-APR-1998) Institute of Genetics, School of Life
            Science, Fudan University, 220 Handan Rd., Shanghai 200433,
            P.R.China
FEATURES             Location/Qualifiers
     source          1. .1387
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="016d06"
                     /tissue_type="brain"
                     /dev_stage="fetus"
     CDS             82. .1134
                     /codon_start=1
                     /product="My018 protein"
                     /protein_id="AAG43131.1"
                     /db_xref="GI:12001976"
                     /translation="MDLKALLSSLNDFASLSFAESWDNVGLLVEPSPPHTVNTLFLTN
                     DLTEEVMEEVLQKKADLILSYHPPIFRPMKRITWNTWKERLVIRALENRVGIYSPHTA
                     YDAAPQGVNNWLAKGLGACTSRPIHPSKAPNYPTEGNHRVEFNVNYTQDLDKVMSAVK
                     GIDGVSVTSFSARTGNEEQTRINLNCTQKALMQVVDFLSRNKQLYQKTEILSLEKPLL
                     LHTGMGRLCTLDESVSLATMIDRIKRHLKLSHIRLALGVGRTLESQVKVVALCAGSGS
                     SVLQGVEADLYLTGEMSHHDTLDAASQGINVILCEHSNTERGFLSDLRDMLDSHLENK
                     INIILSETDRDPLQVV"
BASE COUNT      378 a    323 c    309 g    377 t
ORIGIN

  Query Match           86.2%;  Score 1338.4;  DB 9;  Length 1387;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1361;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy        188 GTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTCCTTCATG 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 GTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTCCTTCATG 84

Qy        248 GATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGCTGAGAGT 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGCTGAGAGT 144

Qy        308 TGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATACACTCTTC 367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 TGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATACACTCTTC 204

Qy        368 CTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGCAGACCTC 427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 CTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGCAGACCTC 264

Qy        428 ATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAACACATGG 487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 ATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAACACATGG 324

Qy        488 AAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCCTCATACA 547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 AAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCCTCATACA 384

Qy        548 GCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGGAGCTTGT 607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 GCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGGAGCTTGT 444

Qy        608 ACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAACCACCGA 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 ACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAACCACCGA 504

Qy        668 GTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGTGAAAGGA 727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 GTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGTGAAAGGA 564

Qy        728 ATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGAACAAACACGG 787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 ATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGAACAAACACGG 624

Qy        788 ATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCTTTCCCGGAAC 847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 ATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCTTTCCCGGAAC 684

Qy        848 AAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCTTCTACATACT 907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 AAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCTTCTACATACT 744

Qy        908 GGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCATGATTGATCGA 967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 GGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCATGATTGATCGA 804

Qy        968 ATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGGGAGAACCTTA 1027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 ATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGGGAGAACCTTA 864

Qy       1028 GAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCTGCAGGGT 1087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        865 GAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCTGCAGGGT 924
```

```
Qy       1088 GTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTTGGATGCTGCT 1147
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 GTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTTGGATGCTGCT 984

Qy       1148 TCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGGCTTTCTTTCT 1207
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        985 TCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGGCTTTCTTTCT 1044

Qy       1208 GACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTATCCTATCAGAG 1267
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1045 GACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTATCCTATCAGAG 1104

Qy       1268 ACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAACACATTCCTA 1327
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1105 ACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAACACATT-CTA 1163

Qy       1328 CAAATCAGCTGGATGCCCAACTTAAATTTGTAACATGAGTCAGTGGGACTGGTGTGCTTC 1387
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1164 CAAATCAGCTGGATG-CCAACTTAAATTTGTAACATGAGTCAGTGGGACTGGTGTGCTTC 1222

Qy       1388 CAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTTATTCACCAAATGTT 1447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1223 CAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTTATTCACCAAATGTT 1282

Qy       1448 CTATCGCTCGTAAGGTAAAACTGTAATATAACTACCATATTAAATAACAAATGTTCATTA 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1283 CTATCGCTCGTAAGGTAAAACTGTAATATAACTACCATATTAAATAACAAATGTTCATTA 1342

Qy       1508 TAAACTCTAGGAAAGATTGAATAAAATCTGTTTACTTAACATTC 1551
             |||||||||||||||||||||||||||||||||||||||||||| |
Db       1343 TAAACTCTAGGAAAGATTGAATAAAATCTGTTTACTTAACATCC 1386
```

```
RESULT 6
BC007654
LOCUS       BC007654                1440 bp    mRNA    linear   PRI 12-JUL-2001
DEFINITION  Homo sapiens, Similar to NIF3 (Ngg1 interacting factor 3, S.pombe
            homolog)-like 1, clone MGC:2595 IMAGE:3346099, mRNA, complete cds.
ACCESSION   BC007654
VERSION     BC007654.1  GI:14043316
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1440)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAY-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site:       http://www.nisc.nih.gov/
            Contact:        nisc_mgc@nhgri.nih.gov
            Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M.,
            Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R.,
            Lim,M., Maduro,Q.L., Masiello,C., Mastrian,S.D., McCloskey,J.C.,
            McDowell,J., Pearson,R., Snyder,B., Stantripop,S., Thomas,P.J.,
            Tiongson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,
            Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 5 Row: i Column: 12
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 10197631.
FEATURES             Location/Qualifiers
     source          1. .1440
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:2595 IMAGE:3346099"
                     /tissue_type="Skin, melanotic melanoma."
                     /clone_lib="NIH_MGC_20"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     CDS             37. .1170
                     /codon_start=1
                     /product="Similar to NIF3 (Ngg1 interacting factor 3,
                     S.pombe homolog)-like 1"
                     /protein_id="AAH07654.1"
                     /db_xref="GI:14043317"
                     /translation="MLSSCVRPVPTTVRFVDSLICNSSRSFMDLKALLSSLNDFASLS
                     FAESWDNVGLLVEPSPPHTVNTLFLTNDLTEEVMEEVLQKKADLILSYHPPIFRPMKR
                     ITWNTWKERLVIRALENRVGIYSPHTAYDAAPQGVNNWLAKGLGACTSRPIHPSKAPN
                     YPTEGNHRVEFNVNYTQDLDKVMSAVKGIDGVSVTSFSARTGNEEQTRINLNCTQKAL
                     MQVVDFLSRNKQLYQKTEILSLEKPLLLHTGMGRLCTLDESVSLATMIDRIKRHLKLS
                     HIRLALGVGRTLESQVKVVALCAGSGSSVLQGVEADLYLTGEMSHHDTLDAASQGINV
                     ILCEHSNTERGFLSDLRDMLDSHLENKINIILSETDRDPLQVV"
BASE COUNT      407 a    329 c    313 g    391 t
ORIGIN
```

```
  Query Match             86.1%;  Score 1336.6;  DB 9;  Length 1440;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1361;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy        187 AGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTCCTTCAT 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 AGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTCCTTCAT 119

Qy        247 GGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGCTGAGAG 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 GGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGCTGAGAG 179

Qy        307 TTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATACACTCTT 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 TTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATACACTCTT 239

Qy        367 CCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGCAGACCT 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 CCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGCAGACCT 299

Qy        427 CATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAACACATG 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 CATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAACACATG 359

Qy        487 GAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCCTCATAC 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCCTCATAC 419

Qy        547 AGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGGAGCTTG 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 AGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGGAGCTTG 479

Qy        607 TACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAACCACCG 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 TACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAACCACCG 539

Qy        667 AGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGTGAAAGG 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 AGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGTGAAAGG 599

Qy        727 AATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGAACAAACACG 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 AATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGAACAAACACG 659
```

```
Qy        787 GATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCTTTCCCGGAA 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 GATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCTTTCCCGGAA 719

Qy        847 CAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCTTCTACATAC 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 CAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCTTCTACATAC 779

Qy        907 TGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCATGATTGATCG 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 TGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCATGATTGATCG 839

Qy        967 AATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGGGAGAACCTT 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 AATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGGGAGAACCTT 899

Qy       1027 AGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCTGCAGGG 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 AGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCTGCAGGG 959

Qy       1087 TGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTTGGATGCTGC 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         960 TGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTTGGATGCTGC 1019

Qy       1147 TTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGGCTTTCTTTC 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1020 TTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGGCTTTCTTTC 1079

Qy       1207 TGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTATCCTATCAGA 1266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1080 TGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTATCCTATCAGA 1139

Qy       1267 GACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAACACATTCCT 1326
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        1140 GACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAACACATT-CT 1198

Qy       1327 ACAAATCAGCTGGATGCCCAACTTAAATTTGTAACATGAGTCAGTGGGACTGGTGTGCTT 1386
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        1199 ACAAATCAGCTGGATG-CCAACTTAAATTTGTAACATGAGTCAGTGGGACTGGTGTGCTT 1257

Qy       1387 CCAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTTATTCACCAAATGT 1446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1258 CCAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTTATTCACCAAATGT 1317

Qy       1447 TCTATCGCTCGTAAGGTAAAACTGTAATATAACTACCATATTAAATAACAAATGTTCATT 1506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1318 TCTATCGCTCGTAAGGTAAAACTGTAATATAACTACCATATTAAATAACAAATGTTCATT 1377

Qy       1507 ATAAACTCTAGGAAAGATTGAATAAAATCTGTTTACTTAACATTCAA 1553
              |||||||||||||||||||||||||||||||||||| |   ||
Db        1378 ATAAACTCTAGGAAAGATTGAATAAAATCTGTTTACTTAAAAAAAAA 1424


RESULT 7
AF283538
LOCUS       AF283538               1425 bp    mRNA    linear   PRI 16-JAN-2001
DEFINITION  Homo sapiens NIF3L1 protein mRNA, complete cds.
ACCESSION   AF283538
VERSION     AF283538.1  GI:12006402
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1425)
  AUTHORS   Tascou,S., Uedelhoven,J., Dixkens,C., Nayernia,K., Engel,W. and
            Burfeind,P.
  TITLE     Isolation and characterization of a novel human gene, NIF3L1, and
            its mouse ortholog, Nif3l1, highly conserved from bacteria to
            mammals
  JOURNAL   Cytogenet. Cell Genet. 90 (3-4), 330-336 (2000)
  MEDLINE   20573864
   PUBMED   11124544
REFERENCE   2  (bases 1 to 1425)
  AUTHORS   Tascou,S., Burfeind,P. and Engel,W.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUN-2000) Institute for Human Genetics, University of
            Goettingen, Heinrich-Dueker Weg 12, Goettingen 37073, Germany
FEATURES             Location/Qualifiers
     source          1. .1425
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
     CDS             47. .1180
                     /codon_start=1
                     /product="NIF3L1 protein"
                     /protein_id="AAG44846.1"
                     /db_xref="GI:12006403"
                     /translation="MLSSCVRPVPTTVRFVDSLICNSSRSFMDLKALLSSLNDFASLS
                     FAESWDNVGLLVEPSPPHTVNTLFLTNDLTEEVMEEVLQKKADLILSYHPPIFRPMKR
                     ITWNTWKERLVIRALENRVGIYSPHTAYDAAPQGVNNWLAKGLGACTSRPIHPSKAPN
                     YPTEGNHRVEFNVNYTQDLDKVMSAVKGIDGVSVTSFSARTGNEEQTRINLNCTQKAL
                     MQVVDFLSRNKQLYQKTEILSLEKPLLLHTGMGRLCTLDESVSLATMIDRIKRHLKLS
                     HIRLALGVGRTLESQVKVVALCAGSGSSVLQGVEADLYLTGEMSHHDTLDAASQGINV
                     ILCEHSNTERGFLSDLRDMLDSHLENKINIILSETDRDPLQVV"
BASE COUNT      385 a    331 c    314 g    395 t
ORIGIN

  Query Match             85.9%;  Score 1334;  DB 9;  Length 1425;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1356;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy        187 AGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTCCTTCAT 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 AGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTCCTTCAT 129

Qy        247 GGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGCTGAGAG 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 GGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGCTGAGAG 189

Qy        307 TTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATACACTCTT 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 TTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATACACTCTT 249

Qy        367 CCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGCAGACCT 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 CCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGCAGACCT 309

Qy        427 CATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAACACATG 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 CATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAACACATG 369

Qy        487 GAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCCTCATAC 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 GAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCCTCATAC 429

Qy        547 AGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGGAGCTTG 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 AGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGGAGCTTG 489

Qy        607 TACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAACCACCG 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 TACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAACCACCG 549

Qy        667 AGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGTGAAAGG 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 AGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGTGAAAGG 609

Qy        727 AATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGAACAAACACG 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        610 AATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGAACAAACACG 669

Qy        787 GATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCTTTCCCGGAA 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 GATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCTTTCCCGGAA 729
```

```
Qy        847 CAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCTTCTACATAC 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 CAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCTTCTACATAC 789

Qy        907 TGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCATGATTGATCG 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 TGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCATGATTGATCG 849

Qy        967 AATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGGGAGAACCTT 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 AATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGGGAGAACCTT 909

Qy       1027 AGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCTGCAGGG 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 AGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCTGCAGGG 969

Qy       1087 TGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTTGGATGCTGC 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 TGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTTGGATGCTGC 1029

Qy       1147 TTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGGCTTTCTTTC 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 TTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGGCTTTCTTTC 1089

Qy       1207 TGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTATCCTATCAGA 1266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1090 TGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTATCCTATCAGA 1149

Qy       1267 GACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAACACATTCCT 1326
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1150 GACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAACACATT-CT 1208

Qy       1327 ACAAATCAGCTGGATGCCCAACTTAAATTTGTAACATGAGTCAGTGGGACTGGTGTGCTT 1386
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db       1209 ACAAATCAGCTGGATG-CCAACTTAAATTTGTAACATGAGTCAGTGGGACTGGTGTGCTT 1267

Qy       1387 CCAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTTATTCACCAAATGT 1446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1268 CCAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTTATTCACCAAATGT 1327

Qy       1447 TCTATCGCTCGTAAGGTAAAACTGTAATATAACTACCATATTAAATAACAAATGTTCATT 1506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1328 TCTATCGCTCGTAAGGTAAAACTGTAATATAACTACCATATTAAATAACAAATGTTCATT 1387

Qy       1507 ATAAACTCTAGGAAAGATTGAATAAAATCTGTTTACTT 1544
              ||||||||||||||||||||||||||||||||||||||
Db       1388 ATAAACTCTAGGAAAGATTGAATAAAATCTGTTTACTT 1425
```

```
RESULT 8
HSM805522
LOCUS       HSM805522              1353 bp    mRNA    linear   PRI 12-JUL-2002
DEFINITION  Homo sapiens mRNA; cDNA DKFZp762L015 (from clone DKFZp762L015).
ACCESSION   AL834430
VERSION     AL834430.1  GI:21740154
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1353)
  AUTHORS   Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B. and
            Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            This clone (DKFZp762L015) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .1353
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp762L015"
                     /tissue_type="melanoma (MeWo cell line)"
                     /clone_lib="762 (synonym: hmel2). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
     polyA_signal    1228. .1233
     polyA_site      1252
BASE COUNT      457 a    283 c    280 g    333 t
ORIGIN
```

```
  Query Match            79.3%;  Score 1232;  DB 9;  Length 1353;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 1254;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy        298 TGCTGAGAGTTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAA 357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGCTGAGAGTTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAA 60

Qy        358 TACACTCTTCCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAA 417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TACACTCTTCCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAA 120

Qy        418 GGCAGACCTCATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTG 477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCAGACCTCATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTG 180

Qy        478 GAACACATGGAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTC 537
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAACACATGGAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTC 240

Qy        538 TCCTCATACAGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCT 597
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCCTCATACAGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCT 300

Qy        598 TGGAGCTTGTACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGG 657
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGGAGCTTGTACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGG 360

Qy        658 AAACCACCGAGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGC 717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAACCACCGAGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGC 420

Qy        718 AGTGAAAGGAATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGA 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGTGAAAGGAATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGA 480

Qy        778 ACAAACACGGATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCT 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACAAACACGGATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCT 540

Qy        838 TTCCCGGAACAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCT 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCCCGGAACAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCT 600

Qy        898 TCTACATACTGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCAT 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCTACATACTGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCAT 660

Qy        958 GATTGATCGAATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGG 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GATTGATCGAATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGG 720

Qy       1018 GAGAACCTTAGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGT 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGAACCTTAGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGT 780
```

```
Qy       1078 TCTGCAGGGTGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTT 1137
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTGCAGGGTGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTT 840

Qy       1138 GGATGCTGCTTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGG 1197
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGATGCTGCTTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGG 900

Qy       1198 CTTTCTTTCTGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTAT 1257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTTCTTTCTGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTAT 960

Qy       1258 CCTATCAGAGACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAA 1317
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCTATCAGAGACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAA 1020

Qy       1318 CACATTCCTACAAATCAGCTGGATGCCCAACTTAAATTTGTAACATGAGTCAGTGGGACT 1377
             |||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1021 CACATT-CTACAAATCAGCTGGATG-CCAACTTAAATTTGTAACATGAGTCAGTGGGACT 1078

Qy       1378 GGTGTGCTTCCAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTTATTC 1437
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 GGTGTGCTTCCAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTTATTC 1138

Qy       1438 ACCAAATGTTCTATCGCTCGTAAGGTAAAACTGTAATATAACTACCATATTAAATAACAA 1497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 ACCAAATGTTCTATCGCTCGTAAGGTAAAACTGTAATATAACTACCATATTAAATAACAA 1198

Qy       1498 ATGTTCATTATAAACTCTAGGAAAGATTGAATAAAATCTGTTTACTTAACATTCAA 1553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 ATGTTCATTATAAACTCTAGGAAAGATTGAATAAAATCTGTTTACTTAACATTCAA 1254


RESULT 9
AX119075
LOCUS       AX119075               1053 bp    DNA     linear   PAT 11-MAY-2001
DEFINITION  Sequence 239 from Patent WO0129221.
ACCESSION   AX119075
VERSION     AX119075.1  GI:14036029
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Conklin,D.C. and Yee,D.P.
  TITLE     Proteins and polynucleotides encoding them
  JOURNAL   Patent: WO 0129221-A 239 26-APR-2001;
            ZymoGenetics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1053
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             1. .1053
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC38611.1"
                     /db_xref="GI:14036030"
                     /translation="MDLKALLSSLNDFASLSFAESWDNVGLLVEPSPPHTVNTLFLTN
                     DLTEEVMEEVLQKKADLILSYHPPIFRPMKRITWNTWKERLVIRALENRVGIYSPHTA
                     YDAAPQGVNNWLAKGLGACTSRPIHPSKAPNYPTEGNHRVEFNVNYTQDLDKVMSAVK
                     GIDGVSVTSFSARTGNEEQTRINLNCTQKALMQVVDFLSRNKQLYQKTEILSLEKPLL
                     LHTGMGRLCTLDESVSLATMIDRIKRHLKLSHIRLALGVGRTLESQVKVVALCAGSGS
                     SVLQGVEADLYLTGEMSHHDTLDAASQGINVILCEHSNTERGFLSDLRDMLDSHLENK
                     INIILSETDRDPLQVV"
BASE COUNT      279 a    253 c    247 g    274 t
ORIGIN

  Query Match           67.8%;  Score 1053;  DB 6;  Length 1053;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-282;
  Matches 1053;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        245 ATGGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGCTGAG 304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGCTGAG 60

Qy        305 AGTTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATACACTC 364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGTTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATACACTC 120

Qy        365 TTCCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGCAGAC 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGCAGAC 180

Qy        425 CTCATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAACACA 484
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTCATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAACACA 240

Qy        485 TGGAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCCTCAT 544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCCTCAT 300

Qy        545 ACAGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGGAGCT 604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACAGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGGAGCT 360

Qy        605 TGTACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAACCAC 664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAACCAC 420

Qy        665 CGAGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGTGAAA 724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGAGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGTGAAA 480

Qy        725 GGAATTGACGGTGTTTCTGTCACTTCTTTTCTGCTAGGACTGGTAATGAGGAACAAACA 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGAATTGACGGTGTTTCTGTCACTTCTTTTCTGCTAGGACTGGTAATGAGGAACAAACA 540

Qy        785 CGGATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCTTTCCCGG 844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGGATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCTTTCCCGG 600

Qy        845 AACAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCTTCTACAT 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AACAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCTTCTACAT 660

Qy        905 ACTGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCATGATTGAT 964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACTGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCATGATTGAT 720

Qy        965 CGAATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGGGAGAACC 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CGAATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGGGAGAACC 780

Qy       1025 TTAGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCTGCAG 1084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTAGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCTGCAG 840

Qy       1085 GGTGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTTGGATGCT 1144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGTGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTTGGATGCT 900

Qy       1145 GCTTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGGCTTTCTT 1204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCTTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGGCTTTCTT 960

Qy       1205 TCTGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTATCCTATCA 1264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTATCCTATCA 1020

Qy       1265 GAGACTGACAGGGACCCTCTTCAGGTGGTATAA 1297
              ||||||||||||||||||||||||||||||||
Db       1021 GAGACTGACAGGGACCCTCTTCAGGTGGTATAA 1053
```

```
RESULT 10
AF284439
LOCUS       AF284439               1836 bp    mRNA    linear   EOD 16-JAN-2001
DEFINITION  Mus musculus NIF3L1 mRNA, complete cds.
ACCESSION   AF284439
VERSION     AF284439.1  GI:12034693
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1836)
  AUTHORS   Tascou,S., Uedelhoven,J., Dixkens,C., Nayernia,K., Engel,W. and
            Burfeind,P.
  TITLE     Isolation and characterization of a novel human gene, NIF3L1, and
            its mouse ortholog, Nif3l1, highly conserved from bacteria to
            mammals
  JOURNAL   Cytogenet. Cell Genet. 90 (3-4), 330-336 (2000)
  MEDLINE   20573864
   PUBMED   11124544
REFERENCE   2  (bases 1 to 1836)
  AUTHORS   Tascou,S., Burfeind,P. and Engel,W.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUL-2000) University of Gottingen, Institute of Human
            Genetics, Heinrich-Duker weg 12, Gottingen 37073, Germany
FEATURES             Location/Qualifiers
     source          1. .1836
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /chromosome="1"
     CDS             508. .1638
                     /codon_start=1
                     /product="NIF3L1"
                     /protein_id="AAG45961.1"
                     /db_xref="GI:12034694"
                     /translation="MLSSAHIVPTSVQRAQSWICRSSRSFMDLKALLSSLNDFASLSF
                     AESWDNVGLLVEPSPPHTVNTLFLTNDLTEEVMDEALQKKADFILSYHPPIFRPMKHI
                     TWKTWKECLVIRALENRVAVYSPHTAYDAAPQGVNSWLAKGLGTCTTRPIHPSRAPDY
                     PTEGAHRLEFSVNRSQDLDKVMSTLRGVGGVSVTSFPARCDGEEQTRISLNCTQKTLM
                     QVLAFLSQDRQLYQKTEILSLEKPLLLHTGMGRLCTLDESVSLAIMIERIKTHLKLSH
                     LRLALGVGRTLESQVKVVALCAGSGGSVLQGVEADLYLTGEMSHHDVLDAASKGINVI
                     LCEHSNTERGFLSELQEMLGVHFENKINIILSETDRDPLRVV"
BASE COUNT      462 a    517 c    450 g    407 t
ORIGIN

  Query Match           55.9%;  Score 868;  DB 10;  Length 1836;
  Best Local Similarity  85.4%;  Pred. No. 6.9e-231;
  Matches  967;  Conservative    0;  Mismatches  165;  Indels    0;  Gaps    0;

Qy        187 AGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTCCTTCAT 246
              |||||| || |  || |    ||   | ||  ||||||||  ||||||| |||||||||
Db        528 AGTCCCTACAAGTGTGCAACGTGCCCAGTCGTGGATCTGCAGGTCTTCCCGTTCCTTCAT 587

Qy        247 GGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGCTGAGAG 306
              |||| |||||||||| || || |||||||||||||||||||||||||| ||||||||||
Db        588 GGATCTGAAGGCTCTTCTCTCGTCCTTGAATGACTTTGCATCCCTCTCATTTGCTGAGAG 647

Qy        307 TTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATACACTCTT 366
              |||||||||| |||||||||||||| |||||||||| |||||||||  |||||||||
Db        648 CTGGGACAATGTGGGATTACTGGTGGAGCCAAGCCCACCCCATACTGTAAATACACTCTT 707

Qy        367 CCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGCAGACCT 426
              |||||||||||||||| ||||| || ||||| ||||  |||||||||| |||||||  |
Db        708 CCTGACCAATGACCTGACGGAGGAGGTCATGGACGAGGCTCTGCAAAAGAAGGCAGACTT 767

Qy        427 CATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAACACATG 486
              |||||||||||||||||| ||||| |||||  |||||||||| |||||  |||| || ||
Db        768 CATTCTCTCCTACCATCCACCTATTTTCCGGCCCATGAAGCACATAACTTGGAAAACCTG 827

Qy        487 GAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCCTCATAC 546
              ||||||| | || |||||||||||||||||||||||||||| | ||||||||| || ||

Db        828 GAAGGAGTGTCTCGTGATCCGGGCTCTGGAGAACAGAGTCGCTGTCTACTCTCCCCACAC 887

Qy        547 AGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGGAGCTTG 606
              |||||||||||| || |||||||| ||||||| |||||||| |||||||||||| ||||
Db        888 AGCCTATGATGCGGCACCCCAGGGAGTCAACAGCTGGTTGGCCAAAGGGCTTGGAACTTG 947

Qy        607 TACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAACCACCG 666
              ||  |||||||||| || ||||||| |||||  ||||||| |||||||||    |||||
Db        948 CACTACCAGGCCCATCCACCCTTCCAGAGCTCCAGACTACCCCACAGAGGGAGCTCACCG 1007

Qy        667 AGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGTGAAAGG 726
              | |||||||||  || |||  || ||||||||||||||||||||||||| || ||| |||
Db       1008 ACTAGAATTCAGTGTGAACCGCAGCCAAGACCTGGACAAAGTCATGTCTACACTGAGAGG 1067

Qy        727 AATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGAACAAACACG 786
              |||  ||||| ||||||||||||||| |||| |||  || |  ||| || ||||| ||
Db       1068 GGTTGGAGGTGTCTCTGTCACTTCTTTTCCTGCCAGGTGTGATGGTGAAGAGCAAACCCG 1127

Qy        787 GATTAATCTGAATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCTTTCCCGGAA 846
              ||| |  |||||||||||||||||  |||||||||||||  |||  |||||||||| | |
Db       1128 GATCAGCCTGAATTGTACTCAGAAGACTTTGATGCAGGTGCTAGCCTTTCTTTCCCAGGA 1187

Qy        847 CAAACAACTTTATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCTTCTACATAC 906
              || |||||||||||||| || |||||||| ||| ||||||||||||||||||| |||||
Db       1188 CAGACAACTTTATCAGAAAACTGAAATTCTTTCATTGGAGAAGCCTTTGCTTCTGCATAC 1247

Qy        907 TGGAATGGGACGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCATGATTGATCG 966
              |||||||||||||||| |||||||||||||||||||||||||||||  ||||| || ||
Db       1248 TGGAATGGGACGGTTGTGCACACTGGATGAATCTGTCTCCCTGGCAATAATGATAGAGCG 1307

Qy        967 AATAAAAAGACACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGGGAGAACCTT 1026
              ||| |||| |||||||||| || || ||| |||||||||| |||||  |||||||||| ||
Db       1308 AATCAAAACACACCTAAAGCTGTCGCATCTTCGCTTAGCTCTTGGAGTGGGGAGAACATT 1367

Qy       1027 AGAGTCTCAAGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCTGCAGGG 1086
              |||||| |||||||||| |||||||||||||||||||||||||| |||| |||||  || ||
Db       1368 AGAGTCCCAAGTCAAAGTTGTGGCCCTGTGTGCTGGTTCTGGGGGCAGTGTTCTACAAGG 1427

Qy       1087 TGTTGAGGCTGACCTTTACCTCACAGGTGAGATGTCCCATCATGATACTTTGGATGCTGC 1146
              || ||||| |||||||||||||||||||| |||||||| ||||||  | ||||||||||
Db       1428 AGTGGAGGCCGACCTTTACCTCACAGGTGAAATGTCCCACCATGATGTTCTGGATGCTGC 1487

Qy       1147 TTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAGGCTTTCTTTC 1206
              |||| |||| || |||||||||| |||||||||||||||||||||||||| ||||||| |||||
Db       1488 TTCCAAAGGGATCAATGTCATCCTTTGTGAACACAGCAACACTGAAAGAGGCTTCCTTTC 1547

Qy       1207 TGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAAGATAAATATTATCCTATCAGA 1266
              ||| |||| ||| ||||||| |  |||||| |||||||||| || |||||||| || ||
Db       1548 TGAGCTTCAAGAAATGCTGGGTGTTCACTTTGAGAATAAGATTAACATTATCCTGTCTGA 1607

Qy       1267 GACTGACAGGGACCCTCTTCAGGTGGTATAATTGCAGAAACATCAGGATAAC 1318
              ||| |||||||||||||| |  ||||| ||| ||||   || ||| |   ||
Db       1608 GACAGACAGGGACCCTCTCCGTGTGGTTTAAATGCAAGGACGTCAAGGACAC 1659


RESULT 11
BD149184
LOCUS       BD149184               796 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD149184
VERSION     BD149184.1  GI:27854942
KEYWORDS    JP 2002191363-A/4027.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 796)
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002191363-A 4027 09-JUL-2002;
            HELIX RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
```

```
            PN   JP 2002191363-A/4027
            PD   09-JUL-2002
            PF   28-JUL-2000 JP 2000280990
            PI   TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
            PI   SAITO,
            PI   JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
            PI   KEIICHI NAGAI,TETSUJI OTSUKI
            PC
            C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
             10,
            PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
             Primer for synthesizing full-length cDNA and use thereof FH   Key
                     Location/Qualifiers
            FT   source          1. .796
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .796
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      192 a    204 c    202 g    195 t      3 others
ORIGIN

  Query Match             46.7%;  Score 725;  DB 6;  Length 796;
  Best Local Similarity   98.4%;  Pred. No. 5.4e-191;
  Matches  762;  Conservative    0;  Mismatches    8;  Indels    4;  Gaps    3;

Qy          1 GTGATTGTTATCTTGGTGCTGCAGAGGACAGCAGAAGAGGAGATTGGGTCAGAAAACTGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         27 GTGATTGTTATCTTGGTGCTGCAGAGGACAGCAGAAGAGGAGATTGGGTCAGAAAACTGC 86

Qy         61 CCTGCCGCACCAGAGCACAGCGCACTAGTGGGACAGGGGTCCTGACTCAGACTTAACTGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 CCTGCCGCACCAGAGCACAGCGCACTAGTGGGACAGGGGTCCTGACTCAGACTTAACTGG 146

Qy        121 CTGTGTCTCGTGGTTTTTCACTGTCCTGGAAAAGGCCTGAAGTGGCACTGAAATGAGGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CTGTGTCTCGTGGTTTTTCACTGTCCTGGAAAAGGCCTGAAGTGGCACTGAAATGAGGCA 206

Qy        181 TAGATGAGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TAGATGAGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTC 266

Qy        241 CTTCATGGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CTTCATGGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGC 326

Qy        301 TGAGAGTTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 TGAGAGTTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATAC 386

Qy        361 ACTCTTCCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 ACTCTTCCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGC 446

Qy        421 AGACCTCATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 AGACCTCATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAA 506

Qy        481 CACATGGAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CACATGGAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCC 566

Qy        541 TCATACAGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGG 600
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        567 TCATACAGCCTATGATGCTGCGCCCCAAGGCGTCAACAACTGGTTGGCTAAAGGGCTTGG 626

Qy        601 AGCTTGTACCTCCAGGCCCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAA 660
              |||||||||| |||||||||||||||||||||||| ||||||||||||||||||||||||
Db        627 AGCTTGTACCTTCAGGCCCATACATCCTTCCAAAGCT-CCAACTACCCTACAGAGGGAAA 685

Qy        661 CCACCGAGTAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGT 720
              ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||| |

Db        686 CCACCGA-TAGAATTCAACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCANT 744

Qy        721 GAAAGGAATTGACGGTGTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGA 774
              ||||||||||||||| |||||||| | |  ||||||||||||||||||||| |
Db        745 GAAAGGAATTGACGGNGTTTCTGT--CCTTCTTTTCTGCTAGGACTGGTAATNA 796


RESULT 12
AL645474/c
LOCUS       AL645474               55374 bp    DNA     linear   PRI 30-JAN-2002
DEFINITION  Human DNA sequence from clone RP5-925F19 on chromosome 1, complete
            sequence.
ACCESSION   AL645474
VERSION     AL645474.10  GI:18476853
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 55374)
  AUTHORS   Van Hellmond,Z.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Feb 1, 2002 this sequence version replaced gi:18135182.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP5-925F19 is from the library RPCI-5 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pCYPAC2
            IMPORTANT: This sequence is not the entire insert of clone
            RP5-925F19 It may be shorter because we sequence overlapping
            sections only once, except for a short overlap.
            The true left end of clone RP11-478P15 is at 53375 in this
            sequence. The true right end of clone RP11-393M18 is at 2000 in
            this sequence.
FEATURES             Location/Qualifiers
     source          1. .55374
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP5-925F19"
                     /clone_lib="RPCI-5"
BASE COUNT    17426 a   9900 c   9927 g  18121 t
ORIGIN

  Query Match             45.6%;  Score 708.4;  DB 9;  Length 55374;
  Best Local Similarity   76.0%;  Pred. No. 4.1e-186;
  Matches 1046;  Conservative    0;  Mismatches  231;  Indels   99;  Gaps   10;
```

```
Qy       198 CAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTCCTTCATGGATTTGAAGG 257
Db      8372 CTGTTCTGCTTAAAGATTCCCTGATCTGCAAATCT---CATTATTTCATGGATTTGAAGG 8316

Qy       258 CTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGCTGAGAGTTGGGACAATG 317
Db      8315 CTCTTCTTGCCTCCTTGAATGACTTTGCATCCCTCTCATTTGCTGAGAGTTGTGACAATA 8256

Qy       318 TTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATACACTCTTCCTGACCAATG 377
Db      8255 TTGGATTACTGGTGGAACCAAGAGCATCACATGCTATAAACACACTCTTCCAGACCAATG 8196

Qy       378 ACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGCAGACCTCATTCTCTCCT 437
Db      8195 ACTTGACTGAGGGAGTGATGGAGGAGGTGCTGTGAAGGAAGGCAGACCTCATTCTCTCCT 8136

Qy       438 ACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAACACATGGAAGGAGCGCC 497
Db      8135 GCCATCCACCAATTTTCAGACCTATAAGGCACATAACCTAGAAAACCTGGAAAGAACTCC 8076

Qy       498 TGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCCTCATACAGCCTATGATG 557
Db      8075 TGGTAATCCAGGGTCTGGAGAACGAAGTTGGTATGTATTTTCCTCACATAGCCTATGATG 8016

Qy       558 CTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGGAGCTTGTACCTCCAGGC 617
Db      8015 CTGCACCCCAGGGAGTTTGCAACTGGTTGGCTAAAGGGCTTGGAGCTTGGACCTCTAGGC 7956

Qy       618 CCATACATCCTTCCAAAGCTCCCAACTACCCTACAGAGGGAAACCACCGAGTAGAATTCA 677
Db      7955 TGATGCATCCTTTCGAAGCTTCCAATGACTCCATAGAG---------------------- 7918

Qy       678 ACGTTAACTACACCCAAGACCTGGACAAAGTCATGTCTGCAGTG-AAAGGAATTGACGGT 736
Db      7917 ------AACCCACCGAAGACCTGGACAAAGTCACGTCTGCAGTGAAAAAGAGTTGTGAAT 7864

Qy       737 GTTTCTGTCACTTCTTTTTCTGCTAGGACTGGTAATGAGGAACAAACACGGATTAATCTG 796
Db      7863 GTTTCTGTCGCTTCTTTGTCTCCTGGGATTGATGATGAAAAACAAACATGACTCAGTCTC 7804

Qy       797 AATTGTACTCAGAAGGCTTTGATGCAGGTGGTAGATTTTCTTTCCCGGAACAAACAACTT 856
Db      7803 AGTTATACTCAGAAGACTTTGATGCAGGTAGTGGCTTTTCTCTCCCAGAGCATAAAATGA 7744

Qy       857 TATCAGAAGACGGAAATTCTGTCACTGGAGAAGCCTTTGCTTCTACATACTGGAATGGGA 916
Db      7743 CATCAGAAGACTGAAATTATGT-ACTGGAGAAGCCTTTGCTTCTACATACTGGAATGGGA 7685

Qy       917 CGGTTATGCACACTGGATGAATCTGTCTCCCTGGCAACCATGATTGATCGAATAAAAAGA 976
Db      7684 TGGTTATGCATACTGGATGAATCTATCTCCCTGACAGCCATGACTGAGTAAATCACAAGC 7625

Qy       977 CACCTAAAACTATCTCATATTCGCTTAGCCCTTGGGGTGGGGAGAACCTTAGAG-TCTCA 1035
Db      7624 CACTTAAAACTATTACATATGTGCTTAGTTCTTAGAGTAGCGAGGACTGAAGAGTTCTAA 7565

Qy      1036 AGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCTGCAGGGTGTTGAGGC 1095
Db      7564 AGTCAAAGTCGTGGCCCTGTGTGCTGGTTCTGGGAGCAGCGTTCTGCAGAGTATAGAAGC 7505

Qy      1096 TGACCTTTACC------------------TCACAGGTGAGATGTCCCATCATGATACTT 1136
Db      7504 TGACCTTTACCTCCAGCTGGGTAAAGGTCATCCCAGGTGGGATGTCTCATCATAATATTC 7445

Qy      1137 TGGATGCTGCTTCCCAAGGAATAAATGTCATCCTCTGTGAACACAGCAACACTGAACGAG 1196
Db      7444 TGAATGCTGCTTCCTAAGGAGTAAATGTCATCCTTTATGAACTCAGCAACACTGAATAAG 7385

Qy      1197 GCTTTCTTTCTGACCTTCGAGATATGCTGGATTCTCACTTGGAGAATAACATAAATATTA 1256
Db      7384 ACTTTCTTTCTGACATTTGAGATATACTGGGTGCTTACTTGAAGAGTAACATTAATATTA 7325

Qy      1257 TCCTATCAGAGACTGACAGGGACCCTCTTCAGGTGGTATAA---TTGCACAAACATCAGG 1313

Db      7324 TCCT-TCAGAGACAAGCAAGGACCTTCTTTATGTAGTATAACGCAGATAGGAACATCAGG 7266

Qy      1314 ATAACACATTCCTACAAATCAGCTGGATGCCCAACTTAAATTTGTAACATGAGTCAGTGG 1373
Db      7265 ATGACACAATC-------------------------------------------TACAAATC 7247

Qy      1374 GACTGGTGTGCTTCCAGAGAGTGTCTTCGAGGGTATCATCATTTCCGGTTTGTTAATCTT 1433
Db      7246 GATTGGGCCCTTTCAGAAGAGTGTCTTAAAATTTATTATAATTTCTGATTTGTTACACTG 7187

Qy      1434 ATTCACCAAATGTTCTATCGCTCGTAAGGT-AAAACTGTAATATAACTACCATATTAAAT 1492
Db      7186 GCTCACCACATGTTTCATAGCTCATAAGGTAAAAACTGAAAAATAACTACCATGTTAAAG 7127

Qy      1493 AACAAATGTTCATTATAAACTCTAGGAAAGATTGAATAAAATCTGTTTACTTAACA 1548
Db      7126 AACAAATATTTATTATAAACTCTACGAAAGACTGACTAAAATCTATTTACTTAAAA 7071
```

RESULT 13
AC037455/c
LOCUS AC037455 170586 bp DNA linear HTG 30-AUG-2001
DEFINITION Homo sapiens chromosome 02 clone RP11-422L5, WORKING DRAFT SEQUENCE, 25 unordered pieces.
ACCESSION AC037455
VERSION AC037455.5 GI:9887641
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 170586)
AUTHORS Smith,D.R.
TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome Sequence Data
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 170586)
AUTHORS Smith,D.R.
TITLE Direct Submission
JOURNAL Submitted (09-APR-2000) Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02453, USA
COMMENT On Aug 24, 2000 this sequence version replaced gi:8569072.
---------- Genome Center
Center: Genome Therapeutics Corperation
Center code: GTC
Web site: http://www.genomecorp.com/
Contact: gtc-seqcenter@genomecorp.com
---------- Project Information
Center project name: hg215
---------- Summary Statistics
Sequencing vector: N/A
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 990315
Consensus quality: 147001 bases at least Q40
Consensus quality: 157396 bases at least Q30
Consensus quality: 159713 bases at least Q20
Insert size: 168303; sum-of-contigs
Quality coverage: 3.8x in Q20 bases; sum-of-contigs
----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1171: contig of 1171 bp in length
* 1172 1271: gap of unknown length
* 1272 2319: contig of 1048 bp in length
* 2320 2419: gap of unknown length
* 2420 3793: contig of 1374 bp in length

```
            *     3794     3893: gap of unknown length
            *     3894     5414: contig of 1521 bp in length
            *     5415     5514: gap of unknown length
            *     5515     8773: contig of 3259 bp in length
            *     8774     8873: gap of unknown length
            *     8874    10904: contig of 2031 bp in length
            *    10905    11004: gap of unknown length
            *    11005    13204: contig of 2200 bp in length
            *    13205    13304: gap of unknown length
            *    13305    16401: contig of 3097 bp in length
            *    16402    16501: gap of unknown length
            *    16502    18227: contig of 1726 bp in length
            *    18228    18327: gap of unknown length
            *    18328    22401: contig of 4074 bp in length
            *    22402    22501: gap of unknown length
            *    22502    25177: contig of 2676 bp in length
            *    25178    25277: gap of unknown length
            *    25278    29892: contig of 4615 bp in length
            *    29893    29992: gap of unknown length
            *    29993    33457: contig of 3465 bp in length
            *    33458    33557: gap of unknown length
            *    33558    37405: contig of 3848 bp in length
            *    37406    37505: gap of unknown length
            *    37506    41734: contig of 4229 bp in length
            *    41735    41834: gap of unknown length
            *    41835    48999: contig of 7165 bp in length
            *    49000    49099: gap of unknown length
            *    49100    56076: contig of 6977 bp in length
            *    56077    56176: gap of unknown length
            *    56177    63604: contig of 7428 bp in length
            *    63605    63704: gap of unknown length
            *    63705    70821: contig of 7117 bp in length
            *    70822    70921: gap of unknown length
            *    70922    80138: contig of 9217 bp in length
            *    80139    80238: gap of unknown length
            *    80239    88780: contig of 8542 bp in length
            *    88781    88880: gap of unknown length
            *    88881   104642: contig of 15762 bp in length
            *   104643   104742: gap of unknown length
            *   104743   124656: contig of 19914 bp in length
            *   124657   124756: gap of unknown length
            *   124757   143891: contig of 19135 bp in length
            *   143892   143991: gap of unknown length
            *   143992   170586: contig of 26595 bp in length.
FEATURES             Location/Qualifiers
     source          1. .170586
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="02"
                     /clone="RP11-422L5"
                     /clone_lib="RPCI-11"
     misc_feature    1. .1171
                     /note="assembly_name:Contig1"
     misc_feature    1272. .2319
                     /note="assembly_name:Contig13"
     misc_feature    2420. .3793
                     /note="assembly_name:Contig17"
     misc_feature    3894. .5414
                     /note="assembly_name:Contig18"
     misc_feature    5515. .8773
                     /note="assembly_name:Contig19"
     misc_feature    8874. .10904
                     /note="assembly_name:Contig20
                     clone_end:T7"
     misc_feature    11005. .13204
                     /note="assembly_name:Contig21"
     misc_feature    13305. .16401
                     /note="assembly_name:Contig22"
     misc_feature    16502. .18227
                     /note="assembly_name:Contig23"
     misc_feature    18328. .22401
                     /note="assembly_name:Contig24"
     misc_feature    22502. .25177
                     /note="assembly_name:Contig25"
     misc_feature    25278. .29892
                     /note="assembly_name:Contig26"
     misc_feature    29993. .33457
                     /note="assembly_name:Contig27"
     misc_feature    33558. .37405
                     /note="assembly_name:Contig28"
     misc_feature    37506. .41734
                     /note="assembly_name:Contig29"
     misc_feature    41835. .48999
                     /note="assembly_name:Contig30"
     misc_feature    49100. .56076
                     /note="assembly_name:Contig31
                     clone_end:SP6"
     misc_feature    56177. .63604
                     /note="assembly_name:Contig32"
     misc_feature    63705. .70821
                     /note="assembly_name:Contig33"
     misc_feature    70922. .80138
                     /note="assembly_name:Contig34"
     misc_feature    80239. .88780
                     /note="assembly_name:Contig35"
     misc_feature    88881. .104642
                     /note="assembly_name:Contig36"
     misc_feature    104743. .124656
                     /note="assembly_name:Contig37"
     misc_feature    124757. .143891
                     /note="assembly_name:Contig38"
     misc_feature    143992. .170586
                     /note="assembly_name:Contig39"
BASE COUNT    50417 a  34173 c  33290 g  50259 t   2447 others
ORIGIN

  Query Match             26.9%;  Score 417;  DB 2;  Length 170586;
  Best Local Similarity   99.5%;  Pred. No. 1.2e-104;
  Matches  439;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy       1111 AGGTGAGATGTCCCATCATGATACTTTGGATGCTGCTTCCCAAGGAATAAATGTCATCCT 1170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128601 AGGTGAGATGTCCCATCATGATACTTTGGATGCTGCTTCCCAAGGAATAAATGTCATCCT 128542

Qy       1171 CTGTGAACACAGCAACACTGAACGAGGCTTTCTTTCTGACCTTCGAGATATGCTGGATTC 1230
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128541 CTGTGAACACAGCAACACTGAACGAGGCTTTCTTTCTGACCTTCGAGATATGCTGGATTC 128482

Qy       1231 TCACTTGGAGAATAAGATAAATATTATCCTATCAGAGACTGACAGGGACCCTCTTCAGGT 1290
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128481 TCACTTGGAGAATAAGATAAATATTATCCTATCAGAGACTGACAGGGACCCTCTTCAGGT 128422

Qy       1291 GGTATAATTGCAGAAACATCAGGATAACACATTCCTACAAATCAGCTGGATGCCCAACTT 1350
             |||||||||||||||||||||||||||||| |||||||||||||||| |||||||||||||
Db     128421 GGTATAATTGCAGAAACATCAGGATAACACATT-CTACAAATCAGCTGGATG-CCAACTT 128364

Qy       1351 AAATTTGTAACATGAGTCAGTGGGACTGGTGTGCTTCCAGAGAGTGTCTTCGAGGGTATC 1410
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128363 AAATTTGTAACATGAGTCAGTGGGACTGGTGTGCTTCCAGAGAGTGTCTTCGAGGGTATC 128304

Qy       1411 ATCATTTCCGGTTTGTTAATCTTATTCACCAAATGTTCTATCGCTCGTAAGGTAAAACTG 1470
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128303 ATCATTTCCGGTTTGTTAATCTTATTCACCAAATGTTCTATCGCTCGTAAGGTAAAACTG 128244

Qy       1471 TAATATAACTACCATATTAAATAACAAATGTTCATTATAAACTCTAGGAAAGATTGAATA 1530
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     128243 TAATATAACTACCATATTAAATAACAAATGTTCATTATAAACTCTAGGAAAGATTGAATA 128184

Qy       1531 AAATCTGTTTACTTAACATTC 1551
             |||||||||||||||||||||
Db     128183 AAATCTGTTTACTTAACATTC 128163


RESULT 14
AC005037/c
```

```
LOCUS       AC005037              190508 bp    DNA     linear   PRI 30-SEP-2000
DEFINITION  Homo sapiens BAC clone RP11-469M7 from 2, complete sequence.
ACCESSION   AC005037
VERSION     AC005037.2  GI:4827310
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 190508)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
   PUBMED   9847074
REFERENCE   2  (bases 1 to 190508)
  AUTHORS   Abbott,A. and Le,T.
  TITLE     The sequence of Homo sapiens BAC clone RP11-469M7
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 190508)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 190508)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAY-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 190508)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-SEP-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On May 14, 1999 this sequence version replaced gi:3309089.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0469M07
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO.  For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8.  The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at the Roswell Park Cancer Institute
            (http://bacpac.med.buffalo.edu)
            VECTOR:  pBACe3.6
            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-13J8; the clone sequenced
            to the right is RP11-91M5.  Actual start of this clone is at base
            position 1 of RP11-469M7; actual end is at 190508 of RP11-469M7.
FEATURES              Location/Qualifiers
     source           1. .190508
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="2"
                      /map="2"
                      /clone="RP11-469M7"
                      /clone_lib="RPCI-11"
     repeat_region    155. .222
                      /rpt_family="L2"
     repeat_region    982. .1299
                      /rpt_family="Alu"
     repeat_region    1502. .1533
                      /rpt_family="(CA)n"
     repeat_region    1535. .1823
                      /rpt_family="Alu"
     repeat_region    2005. .2308
                      /rpt_family="Alu"
     repeat_region    2392. .2562
                      /rpt_family="L1"
     repeat_region    3284. .3315
                      /rpt_family="Alu"
     repeat_region    3316. .3338
                      /rpt_family="(CAAA)n"
     repeat_region    3339. .3621
                      /rpt_family="Alu"
     repeat_region    3650. .3743
                      /rpt_family="AT_rich"
     repeat_region    4353. .4664
                      /rpt_family="Alu"
     repeat_region    4727. .5018
                      /rpt_family="Alu"
     repeat_region    5545. .5585
                      /rpt_family="(CATA)n"
     repeat_region    5586. .5873
                      /rpt_family="Alu"
     repeat_region    5876. .6007
                      /rpt_family="Alu"
     repeat_region    6327. .6713
                      /rpt_family="MER21_g"
     repeat_region    6924. .7294
                      /rpt_family="MER2_type"
     repeat_region    8027. .8049
                      /rpt_family="AT_rich"
     repeat_region    8050. .8354
                      /rpt_family="Alu"
     repeat_region    8355. .8661
                      /rpt_family="Alu"
     repeat_region    8356. .8684
                      /rpt_family="AT_rich"
     repeat_region    9450. .9752
                      /rpt_family="Alu"
     repeat_region    10371. .10680
                      /rpt_family="Alu"
     repeat_region    10762. .11038
                      /rpt_family="Alu"
     repeat_region    11039. .11109
                      /rpt_family="Alu"
     repeat_region    11138. .11410
                      /rpt_family="Alu"
     repeat_region    11446. .11663
                      /rpt_family="L1"
     repeat_region    11664. .11959
                      /rpt_family="Alu"
     repeat_region    11960. .12189
                      /rpt_family="L1"
```

```
     repeat_region   12190. .12334
                     /rpt_family="Alu"
     repeat_region   12335. .12360
                     /rpt_family="L1"
     repeat_region   12464. .12514
                     /rpt_family="L1"
     repeat_region   12515. .12822
                     /rpt_family="Alu"
     repeat_region   12823. .12862
                     /rpt_family="L1"
     repeat_region   12857. .13096
                     /rpt_family="L1"
     repeat_region   13101. .13330
                     /rpt_family="MaLR"
     repeat_region   13331. .13634
                     /rpt_family="Alu"
     repeat_region   13635. .13692
                     /rpt_family="MaLR"
     repeat_region   13704. .13990
                     /rpt_family="L1"
     repeat_region   14082. .14387
                     /rpt_family="Alu"
     repeat_region   14402. .14704
                     /rpt_family="Alu"
     repeat_region   14797. .15095
                     /rpt_family="Alu"
     repeat_region   15162. .15196
                     /rpt_family="Mariner"
     repeat_region   15606. .15681
                     /rpt_family="L1"
     repeat_region   15691. .15827
                     /rpt_family="MER1_type"
     repeat_region   15861. .16189
                     /rpt_family="L2"
     repeat_region   16426. .16446
                     /rpt_family="(CAAT)n"
     repeat_region   16447. .16730
                     /rpt_family="Alu"
     repeat_region   17072. .17179
                     /rpt_family="L2"
     repeat_region   17212. .17504
                     /rpt_family="Alu"
     repeat_region   17869. .18139
                     /rpt_family="Alu"
     repeat_region   18168. .18464
                     /rpt_family="Alu"
     repeat_region   18606. .19155
                     /rpt_family="MaLR"
     repeat_region   19454. .19767
                     /rpt_family="Alu"
     repeat_region   19807. .19949
                     /rpt_family="MER1_type"
     repeat_region   19950. .20247
                     /rpt_family="Alu"

  Query Match             26.9%;  Score 417;  DB 9;  Length 190508;
  Best Local Similarity   99.5%;  Pred. No. 1.2e-104;
  Matches  439;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy       1111 AGGTGAGATGTCCCATCATGATACTTTGGATGCTGCTTCCCAAGGAATAAATGTCATCCT 1170
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     110994 AGGTGAGATGTCCCATCATGATACTTTGGATGCTGCTTCCCAAGGAATAAATGTCATCCT 110935

Qy       1171 CTGTGAACACAGCAACACTGAACGAGGCTTTCTTTCTGACCTTCGAGATATGCTGGATTC 1230
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     110934 CTGTGAACACAGCAACACTGAACGAGGCTTTCTTTCTGACCTTCGAGATATGCTGGATTC 110875

Qy       1231 TCACTTGGAGAATAAGATAAATATTATCCTATCAGAGACTGACAGGGACCCTCTTCAGGT 1290
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     110874 TCACTTGGAGAATAAGATAAATATTATCCTATCAGAGACTGACAGGGACCCTCTTCAGGT 110815

Qy       1291 GGTATAATTGCAGAAACATCAGGATAACACATTCCTACAAATCAGCTGGATGCCCAACTT 1350
              |||||||||||||||||||||||||||||| |||||||||||||||| |||||||

Db     110814 GGTATAATTGCAGAAACATCAGGATAACACATT-CTACAAATCAGCTGGATG-CCAACTT 110757

Qy       1351 AAATTTGTAACATGAGTCAGTGGGACTGGTGTGCTTCCAGAGAGTGTCTTCGAGGGTATC 1410
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     110756 AAATTTGTAACATGAGTCAGTGGGACTGGTGTGCTTCCAGAGAGTGTCTTCGAGGGTATC 110697

Qy       1411 ATCATTTCCGGTTTGTTAATCTTATTCACCAAATGTTCTATCGCTCGTAAGGTAAAACTG 1470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     110696 ATCATTTCCGGTTTGTTAATCTTATTCACCAAATGTTCTATCGCTCGTAAGGTAAAACTG 110637

Qy       1471 TAATATAACTACCATATTAAATAACAAATGTTCATTATAAACTCTAGGAAAGATTGAATA 1530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     110636 TAATATAACTACCATATTAAATAACAAATGTTCATTATAAACTCTAGGAAAGATTGAATA 110577

Qy       1531 AAATCTGTTTACTTAACATTC 1551
              |||||||||||||||||||||
Db     110576 AAATCTGTTTACTTAACATTC 110556


RESULT 15
AC093681
LOCUS       AC093681              198250 bp    DNA     linear   HTG 09-OCT-2001
DEFINITION  Homo sapiens chromosome 4 clone RP11-663N2, *** SEQUENCING IN
            PROGRESS ***, 26 unordered pieces.
ACCESSION   AC093681
VERSION     AC093681.2  GI:15987253
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 198250)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 198250)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Oct 9, 2001 this sequence version replaced gi:15487524.

            ------------------- Genome Center --------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            Contact: submissions@watson.wustl.edu
            ---------------- Project Information -----------------
            Center project name: H_NH0663N02
            ------------------------------------------------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 26 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1912: contig of 1912 bp in length
            *     1913     2012: gap of unknown length
            *     2013     3539: contig of 1527 bp in length
            *     3540     3639: gap of unknown length
            *     3640     6127: contig of 2488 bp in length
            *     6128     6227: gap of unknown length
            *     6228     9244: contig of 3017 bp in length
            *     9245     9344: gap of unknown length
            *     9345    13288: contig of 3944 bp in length
            *    13289    13388: gap of unknown length
            *    13389    16301: contig of 2913 bp in length
            *    16302    16401: gap of unknown length
            *    16402    20096: contig of 3695 bp in length
            *    20097    20196: gap of unknown length
```

```
            *    20197    24467: contig of 4271 bp in length
            *    24468    24567: gap of unknown length
            *    24568    28680: contig of 4113 bp in length
            *    28681    28780: gap of unknown length
            *    28781    32494: contig of 3714 bp in length
            *    32495    32594: gap of unknown length
            *    32595    35626: contig of 3032 bp in length
            *    35627    35726: gap of unknown length
            *    35727    40050: contig of 4324 bp in length
            *    40051    40150: gap of unknown length
            *    40151    43881: contig of 3731 bp in length
            *    43882    43981: gap of unknown length
            *    43982    50194: contig of 6213 bp in length
            *    50195    50294: gap of unknown length
            *    50295    57454: contig of 7160 bp in length
            *    57455    57554: gap of unknown length
            *    57555    64660: contig of 7106 bp in length
            *    64661    64760: gap of unknown length
            *    64761    72159: contig of 7399 bp in length
            *    72160    72259: gap of unknown length
            *    72260    79835: contig of 7576 bp in length
            *    79836    79935: gap of unknown length
            *    79936    90863: contig of 10928 bp in length
            *    90864    90963: gap of unknown length
            *    90964    99400: contig of 8437 bp in length
            *    99401    99500: gap of unknown length
            *    99501   114190: contig of 14690 bp in length
            *   114191   114290: gap of unknown length
            *   114291   127822: contig of 13532 bp in length
            *   127823   127922: gap of unknown length
            *   127923   143474: contig of 15552 bp in length
            *   143475   143574: gap of unknown length
            *   143575   170444: contig of 26870 bp in length
            *   170445   170544: gap of unknown length
            *   170545   195651: contig of 25107 bp in length
            *   195652   195751: gap of unknown length
            *   195752   198250: contig of 2499 bp in length.
FEATURES             Location/Qualifiers
     source          1. .198250
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /clone="RP11-663N2"
     misc_feature    1. .1912
                     /note="assembly_name:Contig10"
     misc_feature    2013. .3539
                     /note="assembly_name:Contig11"
     misc_feature    3640. .6127
                     /note="assembly_name:Contig12"
     misc_feature    6228. .9244
                     /note="assembly_name:Contig13"
     misc_feature    9345. .13288
                     /note="assembly_name:Contig14"
     misc_feature    13389. .16301
                     /note="assembly_name:Contig15"
     misc_feature    16402. .20096
                     /note="assembly_name:Contig16"
     misc_feature    20197. .24467
                     /note="assembly_name:Contig17"
     misc_feature    24568. .28680
                     /note="assembly_name:Contig18"
     misc_feature    28781. .32494
                     /note="assembly_name:Contig19"
     misc_feature    32595. .35626
                     /note="assembly_name:Contig20"
     misc_feature    35727. .40050
                     /note="assembly_name:Contig21"
     misc_feature    40151. .43881
                     /note="assembly_name:Contig22"
     misc_feature    43982. .50194
                     /note="assembly_name:Contig23"
     misc_feature    50295. .57454
                     /note="assembly_name:Contig24"
     misc_feature    57555. .64660
                     /note="assembly_name:Contig25"
     misc_feature    64761. .72159
                     /note="assembly_name:Contig26"
     misc_feature    72260. .79835
                     /note="assembly_name:Contig27"
     misc_feature    79936. .90863
                     /note="assembly_name:Contig28"
     misc_feature    90964. .99400
                     /note="assembly_name:Contig29"
     misc_feature    99501. .114190
                     /note="assembly_name:Contig30"
     misc_feature    114291. .127822
                     /note="assembly_name:Contig31"
     misc_feature    127923. .143474
                     /note="assembly_name:Contig32"
     misc_feature    143575. .170444
                     /note="assembly_name:Contig33"
     misc_feature    170545. .195651
                     /note="assembly_name:Contig34"
     misc_feature    195752. .198250
                     /note="assembly_name:Contig9"
BASE COUNT    56011 a  42355 c  41444 g  55928 t   2512 others
ORIGIN

  Query Match             26.7%;  Score 414.4;  DB 2;  Length 198250;
  Best Local Similarity   99.8%;  Pred. No. 6.5e-104;
  Matches  415;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        187 AGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTCCTTCAT 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14089 AGTCCCCACGACAGTCCGGTTTGTAGATTCCCTGATCTGCAATTCTTCCCGTTCCTTCAT 14148

Qy        247 GGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGCTGAGAG 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14149 GGATTTGAAGGCTCTCCTTTCTTCCTTGAATGACTTTGCATCCCTCTCGTTTGCTGAGAG 14208

Qy        307 TTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATACACTCTT 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14209 TTGGGACAATGTTGGATTACTGGTGGAACCAAGCCCACCACATACTGTAAATACACTCTT 14268

Qy        367 CCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGCAGACCT 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14269 CCTGACCAATGACCTGACTGAGGAAGTGATGGAGGAGGTGCTGCAAAAGAAGGCAGACCT 14328

Qy        427 CATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAACACATG 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14329 CATTCTCTCCTACCATCCGCCTATCTTCCGACCCATGAAGCGCATAACCTGGAACACATG 14388

Qy        487 GAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCCTCATAC 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14389 GAAGGAGCGCCTGGTGATCCGGGCTCTGGAGAACAGAGTCGGTATCTACTCTCCTCATAC 14448

Qy        547 AGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGGAG 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      14449 AGCCTATGATGCTGCGCCCCAGGGCGTCAACAACTGGTTGGCTAAAGGGCTTGGTG 14504


Search completed: August 23, 2003, 13:41:33
Job time : 3966 secs
```